Geophysical Research Abstracts
Vol. 13, EGU2011-12314, 2011
EGU General Assembly 2011

GeoCENS: Capturing the Long Tail of the Sensor Web

Steve Liang, Shawn Chen, Alec Huang, Leah Li, Rohana Rezel, and David Chang
GeoSensor Web Lab, Department of Geomatics Engineering, University of Calgary, Canada (steve.liang@ucalgary.ca)

In recent years, the number of local scale sensor networks, deployed and operated by individual scientists, is growing exponentially in recent years. The increased amount of available data is being driven by sensor motes which are monitoring changes in everything from climate to water to biological species. With the rapidly increasing number of large-scale sensor network deployments, the vision of a World-Wide Sensor Web (WSW) is becoming a reality. However, these sensors/data collected by individual scientists are not accessible to most other researchers. This class of sensor data can be termed "Dark Data". These data are not indexed or stored at the centralized sensor data portal, so they are nearly invisible to most scientists and other potential users. These sensors/data are much more difficult to find and much less frequently reused. As a result, they are much more likely to remain underutilized and eventually lost. Collectively, these "Dark Sensors" form what is termed the "long tail" of the sensor web.

Building a system to capture the sensor web long tail needs to overcome the following challenges. (1) Handling the heterogeneity: The head of sensor web consists of few organizations generating huge amount of datasets. Therefore, they tend to be much more homogeneous, and this uniformity makes the computing and data management issues simpler. On the other hand, the tail consists of large number of individual sensor publishers. As a result, the tail is much more heterogeneous (e.g., proprietary data formats or different sensor hardware). How to design a system allowing users to access the tail's heterogeneous sensor resources in a coherent and homogeneous manner is the first major challenge; (2) protecting data ownership: It's common for researchers to spend more than 50% of their time and funding on data collection. Without proper mechanisms to protect sensor/data ownership, scientists will not publish and share their sensors and data; (3) motivating scientists to publish sensors/data: There should be incentives to motivate scientists to contribute their sensors/data to the system. The system should be able to create a network effect such that the more high quality sensors/data an individual contribute, more and more value the individual will gain, and, as a consequence of the network effect, more sensor owners will be attracted to publish and share their sensing resources; (4) creating a rich user experience: The number of web content publishers as well as the amount of web content grew exponentially when the intuitive and easy-to-use web publishing tools became available (e.g., blogs, wikipedia, and twitter). Similarly, in order to capture the sensor web long tail, the system should provide an intuitive and coherent user interface allowing users to search for, browse, preview, download and publish sensors/data.

At the GeoSensor Web Lab at the University of Calgary, we are designing an architecture and building a crowd sourcing-based sensor web platform called GeoCENS (Geospatial Cyberinfrastructure for Environmental Sensing). Our aim is to propose innovative approaches and provide the missing software components for capturing the currently missing sensor web long tail. GeoCENS proposes and implements the following solutions in order to address the aforementioned challenges: (1) we designed and implemented an Online Social Network-based (OSN-based) sensor web architecture that exploits sensor web users' collective intelligence. For example, its underlying social graphs, the structure of user interactions, and the users' profiles/preferences allows us to develop a sensor web search engine based on the modern web search algorithms (e.g., Google's PageRank); (2) Recommendation engines are key components in existing non- geospatial long tail systems such as Pandora.com and Rhapsody.com. With the GeoCENS social network infrastructure, we are able to design and develop a sensor web recommendation engine (i.e. a geospatial folksonomy and a collaborative tagging system) that recommends sensors and datasets according to a user's geographical area of interest; (3) we uses the OGC sensor web framework to handle the sensor web heterogeneity and facilitate sensor web interoperability; and (4) we develop an intuitive 3D sensor web browser allowing users to maneuver a 3D virtual globe, in order to browse, visualize, access and publish heterogeneous sensing resources and other relevant information. To our best knowledge, GeoCENS is the first OSN-based sensor web platform, the first sensor web recommendation engine as well as the first virtual globe-based sensor web browser. Currently GeoCENS is at the public beta testing phase, and the system can be accessed at http://www.geocens.ca